

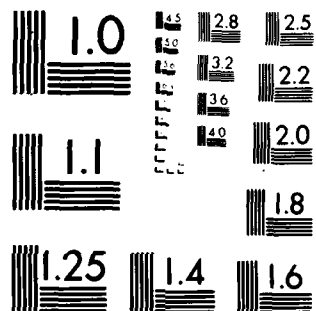
AD-A093 437

DEFENSE MAPPING AGENCY HYDROGRAPHIC/ TOPOGRAPHIC CENT--ETC F/6 5/2
AN ALGORITHM FOR LANGUAGE-INDEPENDENT TRANSFER OF HYDROGRAPHIC --ETC(U)
1980 W G MOIR

UNCLASSIFIED

NL

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOS Eighth Hydrographic Survey Conference	2. GOVT ACCESSION NO. AD A093437	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Algorithm for Language-Independent Transfer of Hydrographic Information.		5. TYPE OF REPORT & PERIOD COVERED
6. AUTHOR(s) William G. Moir		6. PERFORMING ORG. REPORT NUMBER
7. PERFORMING ORGANIZATION NAME AND ADDRESS Defense Mapping Agency Hydrographic/ Topographic Center Washington, D.C. 20315		8. CONTRACT OR GRANT NUMBER(s) 11 1980
8. CONTROLLING OFFICE NAME AND ADDRESS Same as above		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
9. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE
12. 34		13. NUMBER OF PAGES 28
15. SECURITY CLASS. (of this report) Unclassified		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Information Science Coding Algorithm Artificial Grammar Hydrographic Message Chart Number One Computer Communication Theory Nautical Chart Cobol Pre-Controlled Vocabulary Chart Symbol		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A method of communication of hydrographic information is presented. Based on the DMAHTC Chart No. 1, a controlled vocabulary is used to create an algorithm by which information can be transferred in a language-independent, computer-compatible code. Using a specific nautical chart, a controlled vocabulary and this coding algorithm as parameters, a method is presented by which all members of the hydrographic community can share vital information such as Notice to Mariners, in a common format without the limitations imposed by a multiplicity of languages and chart formats.		

LEVEL 1

RECEIVED
JAN 5 1981
C

DD FORM 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408839

LW 81

02 008

AD A093437

FILE COPY

DTIC
COLLECTED
JAN 5 1991

An Algorithm for Language-Independent Transfer
of Hydrographic Information

8th Annual NOS Conference

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution unlimited

Submitted by:
William G. Moir
DMAHTC

4076

Introduction

Since earliest days, man has tried to exploit the 70 percent of the Earth's surface covered by water. As soon as the first intrepid sailor ventured from the coast of Egypt, Phoenicia, or Mesopotamia, he discovered the need for a scientific study of the oceans, a study which continues to this day. Great strides have been made in collecting, evaluating, and verifying hydrographic information in the intervening centuries, particularly since the Challenger Expedition in 1876. It is now within the realm of technical possibility to sound any depth and determine any position with excellent accuracy, and to reshape the shoreline, and even to harness the currents.

Of essence in the advancement of all science, including hydrography, is the communication of information. Assume for a moment that the Vikings did, in fact, visit North America in the late 10th or early 11th century. A momentous discovery to be sure, but one which, in the absence of any clear proof, will always be qualified. Leif Ericsson, reportedly visited Newfoundland, and the Vikings might possibly have reached New England, yet Christopher Columbus gets all the real credit for discovering America. Why? Because his discoveries, even if they were not original, were communicated effectively throughout the known world. Without an effective system of communication any advance in technology or the sciences can spread only very slowly, if at all.

Accession	NTIS
DTIC	Unannounced
Justification	
By	
Distribution	
Availability	
Dist	
A	

So it is with hydrographic and navigational discoveries. During the 16th and 17th centuries only Spain and Portugal had the information necessary to safely navigate to the Orient and back. Contained in books, much like our present day Sailing Directions, these "rutters", as they were called, were prized more highly than gold. The nation which controlled the routes controlled the trade and subsequently accumulated vast wealth.

Since then a spirit of cooperation has emerged out of a concern for the common welfare of mankind and concern for the lives and property involved in oceanic transit. Hydrographic information is no longer a state secret. Cooperation has allowed all maritime nations to conduct their international trade with a high degree of safety.

At issue is whether communication has kept pace with this expansion. Recent years have seen the advent of instantaneous worldwide communication by satellite. Computers are able to handle communications at speeds unimaginable a few years ago. Micrographics, lasers, and fiber optics all promise advances in the near future. But these technologies are not being fully exploited. Consider the Notice to Mariners, which will be the chief example of this report. At present, each of the various maritime nations produces a Notice which differs in format and, sometimes, language from all others. Each is printed

on paper and distributed by mail. To be of use by another hydrographic activity each must be laboriously translated and evaluated.

What is proposed here is not intended to modify this existing system. Instead it is an attempt to use the Notice to Mariners as an example of how the fundamentals of Information Science can be applied to Hydrography. The merger of these disciplines is shown to provide a theoretical basis from which methods may be developed that will be useful in the communication of hydrographic information both nationally and internationally.

The system is designed to be flexible enough to allow a broad spectrum of the Hydrographic community to participate. The procedures will differ slightly, but information can be reported by a cooperating observer aboard a Naval or merchant ship, a member of a survey team, or an information specialist at a national hydrographic activity.

Information Science

It may be beneficial to briefly review the key elements of Information Science. Information Science is a relatively new discipline combining elements of: Computer Science, Library Science, Classification Theory, and Communication Theory. It attempts to define ways in which information can be defined, communicated, stored, retrieved, and otherwise manipulated.

Information is any fact, code, symbol, statistic, or message which has inherent meaning. Information is not to be confused with data; a data item or datum is, in itself, of little meaning. It must be interpreted and evaluated to be considered information.

An example may be useful to clarify the difference. Assume that you are driving a car as it approaches an intersection. On the right shoulder is an eight-sided roadsign, painted red with white lettering. Data elements are: sign, octagonal, red background, and white lettering; collectively these data are sufficient to convey meaning and are thus said to be information. The information is sufficient to trigger a response to the brain initiating a set of actions which stop the car.

Conversely, consider another example. Data elements are: sign, rectangular, white background, and black lettering. No response is triggered. The sign has no information content

per se. It might indicate that the speed limit is X, that the fine for littering is Y, that the highway number is Z or so on. You must read the lettering printed there and interpret it for meaning and relevance to your present needs. If the data on the sign doesn't have meaning for you or if it doesn't reduce uncertainty, no information has been transferred.

Three points which are central to this discussion are suggested by the second road sign. First, the shape and color of the sign are insufficient to convey meaning. It requires additional data for the sign to be useful. This is an important concept to which we will return. There is in every message a minimum number of symbols which are necessary to convey meaning. The system presented in this report attempts to reduce a hydrographic message to only this essential information.

Second, assume that the printing on the sign was incomprehensible or obliterated. Aside from providing one more distraction, the information the sign was intended to convey has been lost. The sign which told of the bridge around the next blind curve being washed out failed to warn of this danger, and thus the driver is unaware of it. And the next driver, and the next.

Like many early explorers these drivers can't benefit from the experience of others who have preceded them. Each driver has to navigate as well as he can. The sign did not communicate the information so there never really was any information.

Finally, suppose that the sign was clearly legible but in a language you did not understand. It carries no information for you. Even though it carried a sufficient, unambiguous message, it conveys no information to someone who doesn't know the code.

Alternatively, consider our stop sign again. The shape, the color, and the code are each sufficient to convey meaning. A white on red sign signals stop. An octagonal sign signals stop. To an English-speaking driver "stop" signals stop. The message is clear, all white on red signs are stop signs, all octagonal signs are stop signs. Even without the lettering, the sign is a symbol which has a clear meaning — STOP!

A key point in our discussion of information is that of uncertainty. Information reduces uncertainty, data does not. Information is unambiguous in form and does not introduce a possibility of misunderstanding or misinterpretation. A symbol in a message which carries an ambiguous meaning reduces the information content of the message.

In order to reduce the ambiguity of symbols in our everyday speech we add modifiers. Within the framework of grammar, modifiers reduce uncertainty and ambiguity. For example, it is possible to speak of a rock in terms of hydrography. As a term, "rock" is ambiguous. It requires modifiers to define it more clearly. It could be a rock above water, a rock which dries at certain stages of the tide (which must be specified), a rock awash at a specified datum, or a sunken rock which may or may not be dangerous as defined by depth and position. The modifiers must be present to define the rock.

Each hydrographic feature requires a specific set of modifiers to give it meaning. We have discussed rocks as an example but the same holds true to other features, such as buoys. Buoys can be a variety of shapes, colors or systems (Cardinal, Lateral, IALA, etc.). Depths, whether dragged or sounded, verified, or suspected, can be in fathoms, feet, meters, or shots of anchor chain. A feature has a specific, finite set of modifiers which it must have and no others. The use of modifiers can be said to be precontrolled.

Another fact must be kept in mind. The use of a natural language, i.e., one in actual spoken use such as English, French, Russian, or Japanese, introduces the nuance of grammar and usage to the words of any message. When a sailor gets ready to "hit

the beach" we all know he's not preparing to box with a sand dune, but a word for word translation into Japanese might indicate this.

Nor is English free from ambiguity even in proper usage. Consider the simple declarative sentence: "She told me that she loved me." Now add the all-purpose modifier "only" anywhere in the sentence. In fact, it can be put in any place in the sentence and make sense although all of the possible sentences have slightly different meanings.

To eliminate confusion which can result from the use of a natural language, we may employ an artificially constructed language with a simple grammar which is rigidly adhered to.

Such an artificial language employs a limited or controlled vocabulary; as mentioned earlier this language is pre-controlled. The vocabulary in this case consists of a number of words and symbols which should be immediately understandable to anyone in the Hydrographic community.

The grammar is controlled first, by limiting the message to 80 alphanumeric characters, this will allow the message or character-string to be transcribed and stored by computer. The reader may be aware that some computer systems will accept up to 120-character-string messages. Until such time

as interactive terminals become widely available, however, an 80 character message should suffice.

The grammar is further controlled by assigning each of the 80 fields to a specific mode (alphabetic, numeric, or blank). Each field can have only certain defined values which will always mean the same thing. The grammar dictates the relative positions of the subject, object and action verbs of the message minimizing the confusion which occurs in natural languages.

Hydrographic Messages

The fundamental unit of this coding algorithm is the hydrographic message. By utilizing a common format, information can be freely communicated and clearly understood by all members of the Hydrographic community. We have examined the theoretical basis for using a controlled vocabulary approach with an artificial grammar, let us now examine in detail the prime elements of the message.

Chart Number-Subject

The first data-group on a message contains seven fields of alphanumeric characters and consists of the chart number. Most nautical chart (National Ocean Survey and Defense Mapping Agency Hydrographic/Topographic Center) numbers are six digits, the seventh field is reserved for use when the reported information falls on a plan or inset. On DMAHTC charts, and on most foreign charts of multiple plans, they are designated by letters A, B, C etc. Insets or undesignated plans must be indicated by thier position on the chart.

The coding rules are as follows:

Fields 1-6 chart number 2 to 6 digits, left-justified and blank-filled.

Field 7 plan designator, alphabetic A, B, C etc. or I if a single inset, or T-top, B-bottom, L-left, R-right if multiple undesignated insets.

For several reasons, the nautical chart was chosen as the base upon which all information will be reported. First, because this is the system currently in use for U.S. Notice to Mariners. Second, because this will allow Naval and merchant mariners to cooperate in reporting information. Survey information may also be reported in this manner, either on an existing chart or with a slight modification of the computer programs on the survey sheet.

Action Verb

The assumption is that any change to be made on a chart can be indicated by a single action verb. The verb is coded by a single numeric field giving a total of 10 possible verbs of which five are presently used at DMAHTC.

The coding is as follows:

1. Add
2. Change
3. Delete
4. Relocate
5. Substitute
6. -Ø Have not yet been assigned.

Chart Symbol-Object

The primary premises of this report are that: (1) it is possible to define any chart symbol by a three-symbol alphanumeric code; and (2) that chart corrections are modifications of the symbols which appear on a chart.

While the latter premise appears self-evident, it is nevertheless necessary. Since we have chosen the chart as the data base, we have reversed the traditional basis of reporting hydrographic information. The usual methodology of reporting utilizes the geographic position as a basis; e.g., the buoy in position N has been moved to position M, or the depth in position X is Y fathoms. This scheme proposes that reported information is best suited to update an existing data base; i.e., a specific nautical chart.

To return to our first premise, the three-symbol codes to be used will be essentially those in Chart No. 1, "Nautical Chart Symbols and Abbreviations." For example, if the reference is to a stranded wreck, the code would be 013. The code references the symbol, not the wreck itself.

In a similar way, any symbol can be coded by the originator of the message and decoded by the recipient. The advantage of the code over plain text is that there can be no

question on the part of the recipient as to which of several classes of rocks or which of a score of buoy symbols was intended.

Decoding the message does require that an additional publication be used but this disadvantage is compensated by the increase in accuracy.

An additional, very important advantage of this system is that it is not language-bound. A non-English speaker can easily encode a message.

Position

Transmission of the geographic position is one part of a typical message which cannot be substantially reduced in length. It will require a maximum of 16 symbols to encode a position and this has been allocated.

The coding is as follows:

Latitude degrees (2), minutes (2), seconds (2) decimal (1)

Longitude degrees (3), minutes (2), seconds (2) decimal (1)

Quadrant (1).

The quadrant of the globe can be identified by a single digit:

1. North and West
2. North and East
3. South and West

4. South and East
5. cross Greenwich Meridian, West first
6. cross Equator North first
7. -Ø not assigned

Quadrant codes 5 and 6 are intended to be used when more than one position is necessary as in the case of an area, a cable or pipeline or a relocated aid when the positions fall in different quadrants.

Even though there are 16 fields allocated for the position, few occasions allow a determination to be made to this precision. Therefore, an adjustment will be allowed. Using an odd number of digits after the degrees will indicate that the last is a decimal; e.g., 653 is read as 65.3 degrees, 65233 is read as 65 degrees 23.3 minutes.

Position is to be left-justified and blank-filled if any sub-unit is not used. In time the system can be modified to accept variable length positions.

It may appear that describing the position in terms of quadrants is an unnecessary complication to replace two alphabetic characters with one digit. The intention is not to go to extremes to shorten the message, but to eliminate its dependence on language.

North, South, East, and West and their abbreviated forms are so familiar to us that we overlook the fact that they are English words. This coding algorithm has been designed to be, so far as possible, entirely free of any natural language so that it can be used effectively by anyone, whether he understands English or not. In fact, if a message is properly coded, it will be impossible to tell which language it was encoded from. This should be welcomed by bilingual countries such as Canada and Belgium, multilingual countries such as India, and by other countries such as the Netherlands and Denmark whose languages are in very limited use outside the country.

We shall now briefly examine some of the other elements which are necessary to the message.

Chart nationality (2 numeric)-since we have limited the chart number to 7 digits, it will be necessary to develop a code for the producer of the referenced chart. Up to 100 countries can be accommodated. This list has yet to be developed, but we can use 00 for DMAHTC and 01 for NOS as an example.

Edition (6 numeric)-indicating the edition number (2) year (2) and month (2) of the chart being used.

Last Notice to Mariners (4 numeric)-notice number (2) and year (2) provides an audit trail to insure that all corrections to this chart have been received.

Chart No. 1 used (1 numeric)-rather than impose the use of the U.S. Chart No. 1, it is proposed that a message can be coded from any of 10 lists of chart symbols. If this system is to be accepted, a new chart symbol pamphlet will have to be produced by the participating country. In the interim, a transmitter must use one of the existing charts listed below. To be sure of being able to decode any message, a receiver would need all of these charts. This puts a burden on the receiver, but since only a national hydrographic activity would find it necessary to collect data from all sources, only they could find it necessary to carry a full complement of Chart No. 1s.

The coding is as follows:

1. I.H.O.
2. British Admiralty
3. Canada
4. France
5. Germany, Federal Republic
6. Japan
7. Spain

8. U.S.A.

9. U.S.S.R.

Ø not assigned

This concludes the listing of elements common to all reports. A brief review shows that we have assigned the first fields as follows:

- 1-7 Chart number
- 8-9 Nationality of chart producer
- 10-15 Chart edition
- 16-19 Last Notice to Mariners
- 20 Action verb
- 21 Chart 1 used for encoding
- 22-25 Chart symbol
- 26 Quadrant
- 27-42 Position
- 43-80 Flags, details, and text.

Flags, details, and text will consume a large portion of some messages and less for others. An in-depth analysis is beyond the scope of this report; additional work has to be done. At the present time, however, we can examine the scope of the problem and identify areas where details must be settled.

First, definitions of the terms are in order. A flag is used here in the data processing sense, certain numbers in certain positions will trigger or flag the computer program to take certain actions. For instance, action verb 4 (Relocate) always indicates that there will be a second geographic position in the message. A chart symbol Q (Soundings) will indicate that the numeric value of the sounding will be included. It is expected that these flags will be written into the computer programs to manipulate the messages. If programs are properly written these flags should be user transparent; that is, the user need not worry about them.

Details are those data which are needed to complete the message such as the second position and the sounding value.

Text is that portion of a message which cannot be coded. The use of quotation marks in the message will indicate that everything within is to be printed as is. Text will include proper names, morse code characteristics of radiobeacons, notes, warnings, and legends not included in the symbol codes. While textual information is, no doubt, important, it will usually be in the language of the chart producer and thus subject to the problems related to natural language.

The following is an overview of details which have been identified. The "If-then" statements should be familiar to any computer programmer and should serve as an example of how the programs can be written.

Action Verbs

If 2 Change, then further details are required

If 4 Relocate, then two positions are required

If 5 Substitute, then two symbols are required

Chart Symbols

If A Coastline, then multiple positions are required

If C Land Features, then multiple positions are required

If D Control point, then elevations may be required

If E Units, then quantity is required

If G Ports and Harbors,

6-11 piers, jetties etc., then two positions are required

12-13, 46, 49-51 areas, then four or more positions are required.

If H Topography, then at least two positions are required

*If K Lights, then K10-13 (position) is required as are characteristic K21-26 and color K61-67.

*If L Buoys, then type 1-33 and color 41-48

*If a message refers to a light or buoy, additional coding is required.

If M Radio and Radar Stations, then characteristic signal and frequency are required.

If O Dangers, then depths may be required.

If P Various Limits, then multiple positions may be required.

If Q Soundings, then values of sounding are required. Unit of Measure will be the same as chart.

If R Depth Contours, then multiple positions are required.

Returning to the coding of the message, starting at the next open field (43), we find our flags as we read from left to right. If the action verb is Relocate, we have identified the chart symbol so we can assign the next 16 fields (43-57) for the second geographic position. In a like manner, if the action verb is Substitute we can assign the next 3 fields (43-45) to the second chart symbol.

Action verbs Add and Delete do not automatically require details; they take their keys from the chart symbol.

The Change verb will usually refer to some existing detail on the chart; e.g., light characteristic, buoy, color, radio-beacon frequency, etc. In this case the reference chart symbol should be the symbol to appear on the chart, and details should begin in field 43.

Lights and buoys are the primary object of the Change verb and are the features least easily coded from the existing Chart No. 1. The coding algorithm below is proposed as a partial solution to this problem.

Knn indicates a light.

First digit indicates type of light.

- 1-. any light symbol
- 2-. light symbol with riprap surrounding
- 3-. aero light
- 4-. light beacon
- 5-. light vessel
- 6-. light with obscured sector(s)
- 7-. range or leading light
- 8-. sector light
- 9-. leading light

Second digit is unassigned at this time. In the event a suitable second key can be devised, it can be implemented at some point in the future. At present this field equals Ø.

The use of any K symbol is a flag for details in the following order:

- (A) Characteristic (n numeric)
- (B) Color (1 numeric)
- (C) Period, in seconds (2 numeric)

(D) Notes (2 numeric)

Characteristics are coded as follows:

- 01. F, fixed
- 02 Fl, flashing
- 03 OCC, occulting
- 04-n unassigned

The (A) Characteristic field is of variable length to allow for group characteristics such as group flashing (3 & 2). Characteristic is left-justified on field 42 and terminated by a single blank.

Color is coded as follows:

- 1. amber or yellow
- 2. blue
- 3. green
- 4. orange
- 5. red
- 6. white
- 7. black
- 8. Ø not assigned

This color coding is also used with buoys and beacons.

Period in seconds is self-explanatory.

Notes are assorted text such as "Exiting" and "destr" which appear under K in Chart No. 1. They are coded the same as any other chart symbol except that they do not take a K in this position.

Light code K80, sectors, sets a flag for an additional group. Sectors are coded: color code (1 numeric), arc (3 numeric, 000°-360°). Any arc not assigned is assumed to be white.

Obscured sector takes the same form; color code, arc, except that color code, if not otherwise noted, is presumed to be an obscured portion.

If any of these details cannot be determined, this can be indicated by leaving a blank.

If any details cannot be coded from the above tables, they can be included as text, for instance light vessel "Borkum Riff" or "Nantucket."

Buoys present a less challenging but, nevertheless, troublesome problem. There are a number of different symbols in use for each type of buoy. If one symbol can be decided upon, then it will eliminate confusion. At present Chart No. 1 lists symbols for 33 different types of buoys, and the list appears

adequate enough to allow coding by this algorithm (Lnn). A buoy does require details, primarily color, but also top-marks and radar reflectors.

As mentioned earlier, buoys can use the same color codes as lights with the exception that two fields must be allotted to allow for the large number of two-color buoys. When, as in the case of a junction or danger buoy, the color of the top band makes a difference, it is given first.

This algorithm is an attempt to formalize all hydrographic messages in a rigid, logical, albeit, somewhat complicated manner. Each code group represents specific information dependent for its meaning on the groups which precede it.

You will note that an area symbol will trigger a flag to require one or more positions in the details section, a depth symbol to require a discrete number and so on. There will be occasions where a message will be longer than 80 characters. In such cases there must be some method of continuing the message. Therefore, we shall introduce the final two codes which are primarily for record-keeping purposes.

Correction number (2 numeric) is to be used primarily by producers of Notice to Mariners, to indicate, for example, that this is the 76th correction to this edition of the Chart.

This code is optional, and if used it is preceded by a non-alphanumeric character such as a comma in field 76.

End of character-string (1 numeric) is always used and field 80 is reserved for it. It is coded as follows:

Ø end of message

1 additional

This concludes my discussion of the coding algorithm itself. The balance of my remarks will address implementation.

Implementation

The coding algorithm may seem, from the text, to be quite complicated. In fact, it isn't. It is no more complicated than the scheme in use for U.S. Notice to Mariners and immeasurably less complicated than plain English.

It will, of course, be necessary to prepare a coding guide for users of this system. Anyone who has ever been a weather observer aboard ship, and those from NOS who are familiar with the data collection program for weather, can visualize just such a guide which has been in common use for many years.

NOAA's weather report collection program enlists the assistance of cooperating weather observers on board ships. I can recall filling out coding sheets many times. Even without

knowing anything about computers, the guide takes the observer through the report in a logical, step-by-step procedure and allows him to encode a great deal of information on a single line on a coding sheet.

At the same time the coding guide is prepared, the Chart No. 1 can be revised eliminating unnecessary symbols and providing each useful symbol with a discrete number.

Once these tasks are accomplished, a test period can take place. Messages can be coded and decoded by hand and compared. If no information is lost or garbled, then the computer programs can be written.

The computer programs need not be very sophisticated. This algorithm was developed with the computer in mind; nothing should present any difficulties for a working programmer, particularly one who is skilled in writing in COBOL.

Conclusion

The chief value of this system, indeed its only real usefulness, lies in its use as a kind of "lingua franca" for the hydrographic community. A century or more ago all diplomacy in Europe was conducted in French. Similarly, English is the universal language of commercial air transport; a French airliner landing in Istanbul communicates with the tower in English. To communicate hydrographic information effectively, a common language is needed. It is to this need for a common language that this report is addressed; such a common language is possible. Whether it is this algorithm, some variation of it, or a completely different scheme is of little consequence as long as the old parochial communication methods are allowed to become extinct.

A modern communication algorithm will allow hydrographic messages to be transmitted by radio, telephone, or telex. Consider, on the other hand, the delays inherent in the present system. A message, in this case a Notice to Mariners, is laboriously written out, edited, typed, set in type, printed, bundled, mailed, sorted at the receiving end, delivered to the proper office (hopefully), and laboriously translated word by word if not sent to a translator (involving most of the same steps again). Is a month to six weeks sufficiently fast communication for information this vital? I think not!

The world is now entering the Information Age, which many have compared to the Industrial Revolution. This is, perhaps, overstating the fact a bit, but the Hydrographic community must exploit the new information technology.